

DO INTERVIEWERS MODERATE THE EFFECT OF MONETARY INCENTIVES ON RESPONSE RATES IN HOUSEHOLD INTERVIEW SURVEYS?

ELIUD KIBUCHI*

PATRICK STURGIS

GABRIELE B. DURRANT

OLGA MASLOVSKAYA

As citizens around the world become ever more reluctant to respond to survey interview requests, incentives are playing an increasingly important role in maintaining response rates. In face-to-face surveys, interviewers are the key conduit of information about the existence and level of any incentive offered and, therefore, potentially moderate the effectiveness with which an incentive translates nonproductive addresses into interviews. Yet, while the existing literature on the effects of incentives on response rates is substantial, little is currently known about the role of interviewers in determining whether or not incentives are effective. In this article, we apply multilevel models to three different face-to-face interview surveys from the United Kingdom, which vary in their sample designs and incentive levels, to assess whether some interviewers are more successful than others in using incentives to leverage cooperation. Additionally, we link the response outcome data to measures of interviewer characteristics to investigate whether interviewer variability on this dimension is systematically related to level of experience and

ELIUD KIBUCHI, PATRICK STURGIS, GABRIELE B. DURRANT, and OLGA MASLOVSKAYA are with ESRC National Centre for Research Methods (NCRM) and Department of Social Statistics and Demography, School of Economic, Social and Political Sciences, University of Southampton, Highfield Campus, Southampton, SO17 1BJ, UK.

This work was supported by the Economic and Social Research Council (ESRC) as part of Work package 1 of the ESRC National Centre for Research Methods (2014–2019), grant number ES/L008351/1; and the Economic and Social Research Council PhD Studentship grant number ES/J500161/1.

*Address correspondence to Eliud Kibuchi, ESRC National Centre for Research Methods (NCRM) and Department of Social Statistics and Demography, School of Economic, Social and Political Sciences, University of Southampton, Highfield Campus, Southampton, SO17 1BJ, UK; E-mail: emk1g15@soton.ac.uk.

demographic characteristics. Our results show significant and substantial variability between interviewers in the effectiveness of monetary incentives on the probability of cooperation across all three surveys. However, none of the interviewer characteristics considered are significantly associated with more or less successful interviewers.

KEYWORDS: Face-to-face; Incentives; Interviewers; Multilevel.

1. INTRODUCTION

It is widely acknowledged that low and declining response rates pose an existential threat to conventional approaches to data collection in survey research (Brick and Williams 2013; Couper 2013; Meyer, Mok, and Sullivan 2015; Miller 2017). In response to this pressing challenge, survey methodologists have invested considerable time and resources into investigating features of the survey process which can be leveraged to increase the probability of cooperation among sampled units (Groves and Couper 1998; Groves and Heeringa 2006). As the primary interface between survey organizations and sample members, interviewers are key to this endeavour (Morton-Williams 1993; Campanelli, Sturgis, and Purdon 1997; West and Blom 2017). They are a key factor in response rates for household interview surveys remaining substantially higher than all other available modes, although they also come at a commensurately higher cost. A large number of studies in a broad range of contexts have now established that demographic, attitudinal, and behavioral differences between interviewers can account for substantial variability in response rates (Hox and de Leeuw 2002; Hansen 2006; Durrant, Groves, and Steele 2010). For example, Campanelli and O'Muircheartaigh (1999) found that more experienced interviewers were more successful at obtaining contact and cooperation due to more effective calling patterns and an ability to tailor the survey request to sample members' motivations and concerns.

In addition to interviewers, monetary incentives of various kinds have played a central role in strategies for maximising survey cooperation (Singer, Groves, and Corning 1999; Singer, Hoewyk, Gebler, Raghunathan, and McGonagle 1999; Singer 2002). Monetary incentives are considered to operate by acting as a replacement for other nonpecuniary motivations for survey participation such as interest in the survey topic, enjoyment of social interaction, or a sense of civic duty (Singer et al. 1999; Groves, Singer, and Corning 2000). A large body of evidence, predominantly based on randomized experiments, has established that monetary incentives exert a small to moderate positive effect on response rates and that larger incentives tend to produce more substantial effects but with diminishing marginal returns (Church 1993; Singer et al. 1999; Cantor, O'Hare, and O'Connor 2008; Singer and Ye 2013).

Given the sustained focus on the role of interviewers and monetary incentives in the existing survey methodological literature, it is surprising that their

potential *joint* influence has seldom been considered. Because interviewers play such a key role in making contact with and persuading sample members to participate, it is plausible that interviewers vary in how effective they are at leveraging incentives to persuade sample members to provide an interview. For example, some interviewers may tailor their doorstep introductions to highlight the availability of a monetary incentive at households that are most likely to be sensitive to them (Campanelli et al. 1997; Groves and Couper 1998). Similarly, interviewers may feel more confident in their doorstep approach when they know an incentive is available, which may positively affect their persuasive efforts (Singer and Ye 2013). This joint influence is our focus in this article. We analyze data from three different face-to-face interview surveys that included a randomized incentive experiment to identify interviewer influences on the effectiveness of monetary incentives in promoting survey cooperation.

The remainder of the article is structured as follows: We first provide short reviews of the respective literatures on how interviewers and monetary incentives influence survey response, before setting out our expectations regarding the moderating effect of interviewers on the effectiveness of incentives. We then describe the three surveys that form the basis of our analysis and the administrative data on interviewers and areas to which they are linked. This is followed by an exposition of our analysis strategy and presentation of our key findings. We conclude with a consideration of the limitations of our study, a discussion of the implications of our findings for improving survey practice and suggestions on how future research in this area might usefully proceed.

1.1 The Effect of Interviewers on Response Rates

Face-to-face surveys consistently achieve higher response rates than those undertaken by self-administration or by telephone, a difference that is largely attributable to the role of interviewers. Interviewers locate and make repeated calls at sampled addresses, thereby keeping noncontacts to a minimum (Campanelli et al. 1997). Having made contact with a household, they undertake a number of additional tasks including respondent selection within households, conveying information about the survey such as the topic, sponsor, likely duration of the interview, and the availability of incentives (Couper and Schlegel 1998). They also often provide accompanying information about the survey in the form of copies of advanced letters (which will not have been read by all sample members), provide reassurance about the bona fides of the survey, and show identity documentation (Groves and Couper 1998; Groves et al. 2000).

Interviewers also persuade reluctant respondents to provide an interview, thereby minimizing refusals. A range of dispositional factors and behavioral styles have been identified as important in determining how successful interviewers are at preventing refusals. These include an ability to maintain an interaction rather than accept a refusal and to tailor their approach on the

doorstep to specific characteristics of sample units by identifying and presenting aspects of the survey that they judge are likely to be positively valued (Morton-Williams 1993; Campanelli et al 1997; Groves and Couper 1998). For example, an interviewer may remark upon the respondent's garden if they perceive that gardening is likely to be a hobby of the householder, or they might highlight the topic of the survey if they judge from the observable characteristics of the sample member that it is likely to be of interest. Studies which have examined the causes of noncontact and refusal have consistently found significant interviewer effects across a range of sample designs and international contexts (Campanelli et al. 1997; Hox and de Leeuw 2002; Durrant and Steele 2009; Durrant et al. 2010). For example, Blom, de Leeuw, and Hox (2011) found interviewer intraclass correlation coefficients of 0.27 for noncontact and 0.08 for cooperation across ten countries in the 2008 European Social Survey.

Existing research has also considered which characteristics of interviewers are important in producing these effects (Blom and Korbmacher 2013). This has found that experienced interviewers tend to be better at tailoring their approaches to household idiosyncrasies and concerns (Groves and Couper 1998; Lemay and Durand 2002). More experienced interviewers, both in terms of experience on the particular survey and of interviewing more generally, have also been found to obtain higher response rates, even though they are often allocated to more difficult areas (Purdon, Campanelli, and Sturgis 1999; West and Blom 2017). Other studies have found that interviewers with higher levels of self-confidence and more positive appraisals of the likelihood of achieving interviews also obtain higher cooperation rates, an effect which is thought to arise from the positive effect of confidence on the quality of doorstep interactions (Singer and Kohnke-Aguirre 1979; Groves and Couper 1998; Hox and de Leeuw 2002). The existing evidence suggests that interviewer skills and experience in recognizing, interpreting, and addressing visual cues and the confidence and self-belief with which interviewers approach the task of obtaining cooperation on the doorstep are the key mechanisms through which interviewers influence individual cooperation decisions.

1.2 Using Incentives to Increase Response Rates

Under the influential “leverage-salience” theory of survey cooperation (Groves et al. 2000), incentives are postulated to work by acting as a replacement for non-financial motivating factors such as engagement in the topic of the survey, enjoyment of social interaction, and a sense of civic or moral obligation. Incentives may also invoke norms of reciprocity, such that respondents feel a sense of obligation to provide an interview when they are offered or receive an incentive before the interview request is made (Dillman, Smyth, and Christian 2009).

The field of survey research has benefited from a wealth of systematic reviews and meta-analyses of the effects of survey incentives which have

yielded a robust set of conclusions. We know from this body of evidence that monetary incentives are more effective in motivating participation than non-monetary incentives such as pens, calendars, diaries, and so on (Church 1993; Cantor et al. 2008; Singer and Ye 2013). It is also well established that prepaid (or unconditional) incentives tend to produce more substantial effects on response rates than those that are promised (or conditional) on completion of the survey (Church 1993; Singer et al. 1999; Cantor et al. 2008; Lavrakas 2008), though it does not follow from this that they are necessarily more cost-effective (Brick, Montaquila, Hagedorn, Roth, and Chapman 2005). It is also apparent from these studies that the effect of incentives is greater for self-completion surveys and surveys that have a low response rate when no incentive is offered, presumably because there is more scope for the incentive to act as a replacement for nonmonetary motivations among a larger pool of potential nonrespondents (Singer et al. 1999; Mercer, Caporaso, Cantor, and Townsend 2015).

Researchers have also established that the magnitude of the effect of incentives on response rates increases with the size of the incentive. For instance, in a meta-analysis of thirty-nine experimental studies, Singer, Groves, and Coming (1999) found that each dollar of incentive paid resulted in one-third of a percentage point increase in response rate compared with the no incentive condition. However, other studies have found that this “dose-response” relationship is curvilinear, with the size of the increase in the response rate declining with additional increases in the value of the monetary incentive (Gelman, Stevens, and Chan 2002; Cantor et al. 2008; Mercer et al. 2015). In sum, the existing evidence demonstrates that monetary incentives have a robust, positive effect on the probability of survey cooperation.

1.3 The Joint Effect of Interviewers and Incentives on Response Rates

We know that interviewers and incentives have a positive influence on response rates, but what of their joint effect? It seems plausible that interviewers might moderate the effect of incentives on cooperation probability for three inter-related reasons. First, interviewers are the primary conduit of information between survey organization and sample members and are, therefore, essential to ensuring that potential respondents are aware that an incentive is available. While most surveys will highlight incentives in an advanced letter, many respondents do not open—let alone read—them (Stoop 2005). Furthermore, it seems reasonable to assume that those who do not read advanced letters—those who are busy and/or uninterested in the survey topic—are also more likely to be susceptible to monetary incentives. Second, interviewers may have more confidence in the likelihood of obtaining an interview when a monetary incentive is offered. This might exert an additional positive effect on cooperation over and above the influence of the incentive on respondents because higher levels of confidence improve the quality of interviewer approaches

(Groves and Couper 1996; Singer et al. 1999; Singer and Ye 2013). Third, interviewers may vary in the extent to which they tailor their doorstep introductions by highlighting the availability of the incentive at addresses where they believe it is likely to be effective. For example, some interviewers might ask sample members whether they received the letter with information about the payment at an early stage of the interaction, while others do not mention it at all.

Existing research, however, offers little in the way of hard evidence on the question of whether or not interviewers moderate the effects of monetary incentives on cooperation probability. An exception is Singer, Hoewyk, and Maher (2000), who investigated the influence of interviewer expectations on the effect of incentives on cooperation rates using data from the Survey of Consumer Attitudes, a telephone survey of the American public. Singer and Maher randomly assigned interviewers and respondents to three groups: in groups one and two, respondents received an advance letter and a \$5 unconditional incentive, while respondents in group three received an advance letter but no incentive. Interviewers in group one were unaware of the incentive, but interviewers in groups two and three were made aware of the incentive level via messages on their computers. Interviewers in groups one and two achieved response rates of 76 percent and 75 percent, respectively, compared with 62 percent for interviewers in group three. Singer, Hoewyk, and Maher (2000) concluded that, although the unconditional incentive boosted response, interviewer expectations about the likely cooperativeness of sample members had no additional effect. Lynn (2001) found similar evidence from a focus group of interviewers that expectations about the likely impact of incentives on cooperation bore little resemblance to actual response outcomes. While these studies support the conclusion that incentives operate primarily or exclusively via their effects on respondents rather than on interviewers, they do not rule out the possibility that interviewers vary in the effectiveness with which they deploy incentives. We turn next to a direct empirical assessment of this question.

2. DATA

We use data from three different United Kingdom face-to-face interview surveys. These are the 2015 National Survey for Wales Field Test (NSW2015), the 2016 National Survey for Wales Incentive Experiment (NSW2016), and wave one of the United Kingdom Household Longitudinal Study Innovation Panel (UKHLS-IP). All three surveys use stratified random sampling, with addresses selected from the Postcode Address File. The two Welsh surveys randomly select one eligible adult (aged sixteen and over), while UKHLS-IP attempts interviews with all eligible adults (aged eighteen and over) in the household. For UKHLS-IP, a cooperating household is defined as one in which at least one eligible adult provided an interview. The NSW2015 randomly allocated 50 percent of addresses to receive no incentive and 50 percent to receive

£10; NSW2016 also used a 50/50 allocation but with a treatment condition of £5 and a control condition of no incentive. The UKHSL-IP randomly allocated one-third of addresses to receive a £5 incentive, one-third of addresses to receive £5 rising to £10 conditional on all household members completing the survey, and one-third to receive £10. For the analyses presented here, the latter two groups are combined. Incentives in all three surveys were offered conditional on completion of the questionnaire, and allocation of addresses to experimental conditions was implemented within interviewer workloads. We use response outcomes before any re-issuing in order to ensure that the random assignment of incentives within interviewers is maintained. However, detailed first issue outcomes were not available for the UKHLS-IP, so we are only able to model response/nonresponse for this survey rather than cooperation conditional on contact. Analysis of the two Welsh surveys shows that the results are substantively the same for both response and cooperation, so we do not consider this to be an important limitation. More detailed information about the design of each survey is provided in the appendix.

Each survey was linked to administrative data held on interviewers by the respective survey agencies. These were age, sex, and experience (number of years working for the agency). We use these variables to assess whether interviewer characteristics are associated with variability in the effectiveness of deploying incentives. For the UKHLS-IP, we also link aggregate census variables from the 2011 census to the sample file. A total of twenty-one census count variables were combined using a factorial ecology model (Rees 1971), with a total of five neighborhood indices extracted. These measures cover the extent of “concentrated disadvantage” (areas with a higher number of single parent families, those on income support and unemployed, fewer people in managerial and professional occupations, and less owner occupiers), “urbanicity” (high population density and domestic properties, and relatively little green space), and “population mobility” (higher levels of in- and out-migration and more single person households). We also account for differences in the neighborhood age structure (with higher scores for areas with a younger population), housing structure (higher scores for areas with more terraced and vacant properties), and the police recorded crime rate. We model both the response rate and the cooperation rate. Response rate is defined based on AAPOR RR2 (AAPOR 2016) as,

$$RR = \frac{(I + P)}{(I + P) + (R + NC + O) + (UE(NC) + UE)},$$

where RR denotes Response Rate, I denotes interview, P denotes partial interviews, R denotes refusals, NC denotes noncontacts, O denotes other unproductive, UE(NC) denotes unknown eligibility (noncontacted), and UE denotes unknown eligibility. The cooperation rate (CR) conditions on those contacted and is defined as

Table 1. Incentives and Fieldwork Outcomes Before Reissuing for the Three Surveys

	NSW2015		NSW2016		UKHLS-IP	
	£10	£0	£5	£0	£10	£5
Interviews	1,387	1,228	1,772	1,664	1,020	469
Refusals	640	670	954	961	—	—
Noncontact	285	289	265	250	—	—
Other nonresponse	285	273	230	233	—	—
Total nonresponse	1210	1232	1,449	1,444	660	374
Ineligible	368	370	383	359	175	88
Cooperation Rate	68%	65%	65%	63%	—	—
Response Rate	53%	50%	55%	54%	61%	56%
Total issued sample	2,965	2,830	3,604	3,467	1,855	931

NOTE.— Only total nonresponse is available for UKHLS-IP at first issue.

$$CR = \frac{(I + P)}{(I + P + R)}.$$

Response outcomes for the three surveys are presented in table 1. The response rates were higher in the incentive condition for all three surveys, with the NSW2015 and NSW2016 having a 3 and 2 percentage points higher cooperation rates for the incentivized households respectively and the UKHLS-IP having a 5 percentage points higher response rate in the incentive condition. The difference is statistically significant at the 95 percent level of confidence (using a χ^2 test) for UKHLS-IP and NSW2015 but not for NSW2016.

Table 1 demonstrates the cooperation rate was higher in the incentivized condition for all three surveys, though the difference was statistically significant in only one. Next, we proceed to a multivariate analysis to assesses whether these average differences in cooperation rates are constant across interviewers or whether some interviewers are more successful at using the incentive to convert refusals into interviews.

3. ANALYSIS

The influence of interviewers on the effectiveness of incentives on survey cooperation and response is assessed using multilevel logistic regression models (Hox and de Leeuw 2002; Durrant and Steele 2009; Goldstein 2010). The model applied here has the following form: Let y_{ij} denote the binary response for household i ($i = 1, \dots, i$), interviewed by interviewer j ($j = 1, \dots, j$) where,

$$y_{ij} = \begin{cases} 1 & \text{cooperation/response} \\ 0 & \text{refusal/nonresponse} \end{cases}.$$

y_{ij} is assumed to follow a Bernoulli distribution, with conditional response probabilities $\pi_{ij} = Pr(y_{ij} = 1)$ and $1 - \pi_{ij} = Pr(y_{ij} = 0)$. The multilevel logistic regression model accounting for interviewer effects takes the form

$$\log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = \beta_0 + \beta_1 x_{1ij} + \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_j\boldsymbol{\alpha} + \mu_{0j} + \mu_{1j},$$

where β_0 is the intercept, β_1 is the coefficient for the incentive condition; x_{1ij} is a dummy indicator of the incentive group for household i within the assignment of interviewer j ; \mathbf{x}'_{ij} is a vector of household-level characteristics with coefficient vector $\boldsymbol{\beta}$; \mathbf{z}'_j is a vector of interviewer-level covariates with coefficient vector $\boldsymbol{\alpha}$; μ_{0j} is a random intercept; and μ_{1j} is a random coefficient for the incentive dummy. The random intercept and slope, μ_{0j} and μ_{1j} , are assumed to follow a normal distribution with zero mean and variance matrix Ω_μ defined as

$$\begin{bmatrix} \mu_{0j} \\ \mu_{1j} \end{bmatrix} \sim N(0, \Omega_\mu) \text{ where } \Omega_\mu = \begin{bmatrix} \sigma_{\mu 0}^2 & 0 \\ \sigma_{\mu 01} & \sigma_{\mu 1}^2 \end{bmatrix},$$

where $\sigma_{\mu 0}^2$ is the intercept variance, $\sigma_{\mu 1}^2$ is the variance in slope, and $\sigma_{\mu 01}$ is the covariance between intercepts and slope. The positive values of $\sigma_{\mu 01}$ indicate that the effect of the incentive is greater for interviewers with higher cooperation/response rates, although negative values indicate the opposite. Cross-level interactions between interviewer characteristic variables and the incentive variable are included to test whether observable characteristics of interviewers are associated with variability in the effectiveness of deploying incentives.

In standard face-to-face survey designs such as those considered here, identification of interviewer effects is complicated by the confounding of interviewer assignments and areas (Campanelli and O'Muircheartaigh 1999; Durrant et al. 2010). Failure to account for differences in the area-level composition of interviewer assignments can result in overestimation of the magnitude of interviewer effects (O'Muircheartaigh and Campanelli 1998). Where there is an overlap between interviewer assignments and areas, this can be mitigated using a cross-classified, multi-level model (Durrant and Steele 2009). However, this could not be done for the three datasets analyzed here because it was not possible to obtain geographic identifiers for the two Welsh surveys, and the UKHLS-IP did not contain sufficient crossing of interviewers and areas to implement a cross-classified model. We therefore control for area characteristics as fixed effects in the models for the UKHLS-IP data and assess the impact this has on the interviewer random effects.

Models are estimated using Markov Chain Monte Carlo (MCMC) methods using MLwiN software (Fearn, Gelman, Carlin, Stern, Rubin et al. 2004; Browne, Kelly, Charlton, and Pillinger 2016). The starting values for the fixed effects are the second-order penalized quasi-likelihood (PQL) estimates. Priors for the variance matrix are assumed to follow an inverse Wishart distribution $p(\Omega_\mu^{-1}) \sim \text{Wishart}_n(n, \cdot)$, where n is the number of rows in the variance matrix and is an estimate for the true value of the variance matrix Ω_μ (Browne et al. 2016). Because we are using MCMC, we also assess significance of coefficient estimates using the change in model deviance information criterion (DIC) (Spiegelhalter, Best, Carlin, and van der Linde 2002). Deviance information criterion balances model fit and model complexity by taking the sum of the posterior expectation (mean) of the deviance function (\bar{D}) and the effective number of parameters (pD). When comparing DIC values, a model with a DIC value of at least 3 points lower than the previous model is considered to have a significantly better fit (Spiegelhalter et al. 2002; Rasbash, Steele, Browne, Goldstein, Charlton et al. 2012). The models had a burn-in length of 10,000 and 200,000 iterations. In order to avoid undue influence of starting values, different burn-in lengths were tried, as recommended by Fearn et al. (2004). The Brooks-Draper and Raftery-Lewis diagnostics were checked to determine how long the chain must be run to obtain accurate posterior estimates (Browne et al. 2016).

4. RESULTS

Table 2 presents the coefficient estimates, their standard deviations, and the corresponding 95 percent credible intervals for the NSW2015 and NSW2016 models. As we saw in table 1, the coefficients for the incentive fixed effect are positive for both surveys, although only for NSW2015 does the 95 percent credible interval not include zero. The random coefficient variances of 0.09 and 0.07 are both significant, indicating that interviewers vary in the effectiveness with which they deploy incentives. The DIC decreases by 8.0 for NSW2015 and by 15.1 for NSW2016 when the interviewer random coefficient is introduced, indicating an improvement in model fit.

The cross-level interactions between the three interviewer characteristic variables—age, sex, and experience—and the incentive dummy are all nonsignificant, indicating that these interviewer characteristics do not explain between-interviewer variability in the effectiveness of incentives on cooperation. The DIC change, when these interaction terms are added, are -2 for NSW2015 and 4.1 for NSW2016, indicating a small improvement in model fit after the inclusion of these interactions for NSW2016. The covariance between the random intercept and random coefficient σ_{u01} is nonsignificant for both surveys, with a point estimate of 0.02 for NSW2015 and of -0.02 for NSW2016. This indicates that the effectiveness of incentive deployment between interviewers

Table 2. Estimated Coefficients for the Final Model for NSW2015 and NSW2016 Cooperation)

Variable	Category	NSW2015				NSW2016			
		β	SD	95% credible interval		β	SD	95 percent Credible Interval	
Intercept		0.239	0.320	-0.409	0.857	0.651	0.100	0.455	0.848
Incentive (no incentive)	Incentive	0.288	0.363	-0.433	0.991	0.114	0.111	-0.106	0.331
Interviewer age (young)	Lower middle	0.048	0.237	-0.406	0.526				
	Upper Middle	0.345	0.233	-0.103	0.808	0.154	0.155	-0.154	0.456
	Old	0.428	0.278	-0.102	0.982				
Interviewer Experience (less)	Lower middle	-0.027	0.255	-0.535	0.471				
	Upper middle	0.403	0.285	-0.170	0.959	-0.309	0.195	-0.691	0.074
	Highest	0.386	0.274	-0.159	0.929	-0.371	0.224	-0.804	0.071
Interviewer Sex (Female)	Male	0.048	0.237	-0.406	0.526	-0.134	0.132	-0.395	0.124
Incentive (£10 per adult)	£10 per adult *Male	0.035	0.180	-0.316	0.385	-0.268	0.147	-0.556	0.019
*Gender (Female)									
Incentive (£10 per adult)	£10* Lower middle	-0.021	0.269	-0.540	0.506				
*Age (young)									
	£10* Upper Middle	-0.124	0.264	-0.638	0.394	0.063	0.172	-0.271	0.402
	£10* Old	-0.412	0.317	-1.035	0.216				
Incentive (£5) *Experience (less)	£10*Lower Middle	-0.027	0.281	-0.587	0.516				
	£10*Upper Middle	-0.176	0.312	-0.791	0.434	0.546	0.222	0.107	0.984
	£10*Highest	0.109	0.297	-0.480	0.688	0.006	0.246	-0.476	0.486
$\sigma^2_{\mu_0} = var(\mu_{0j})$		0.199	0.065	0.097	0.349	0.137	0.045	0.067	0.243
$\sigma^2_{\mu_1} = var(\mu_{1j})$		0.085	0.047	0.025	0.205	0.070	0.037	0.022	0.161
$\sigma_{\mu_01} = cov(\mu_{0j}, \mu_{1j})$		0.020	0.042	-0.073	0.096	-0.028	0.035	-0.108	0.016
DIC		4765.957		6567.170					

N = 3, 925 (for NSW 2015) N = 5, 175 (for NSW 2016).

NOTE.—NSW2015: £10 Incentive

NSW2016: £5 Incentive

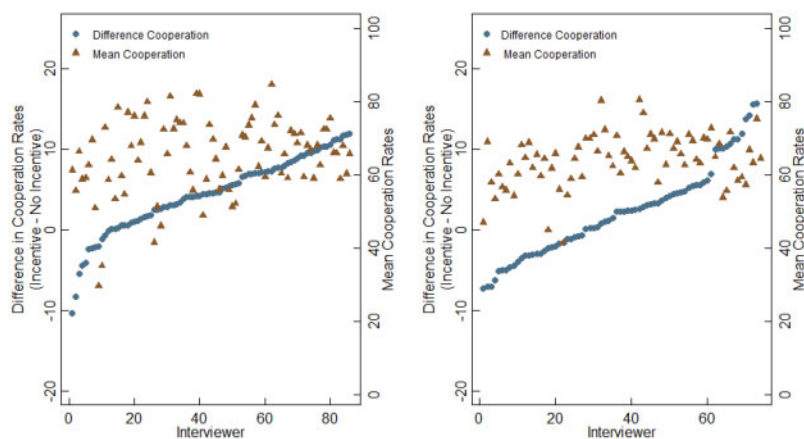


Figure 1. Difference in Predicted Rates of Cooperation for Incentive and Nonincentive Households by Interviewer for NSW 2015 (Left Panel) and NSW 2016 (Right Panel).

is not related to the overall cooperation rate an interviewer achieves on their assignment of addresses.

Figure 1 plots the difference in the mean predicted rates of cooperation for each interviewer derived as fitted values from the models in table 2. Each dot in figure 1 represents an interviewer, with the left Y axis being the difference in the cooperation rates for households in the incentive and nonincentive conditions. The triangles show the mean overall cooperation rate (plotted against the right Y axis) for each interviewer across all eligible households in their assignment. There is substantial variability across interviewers in the effectiveness of the incentive in obtaining cooperation. For NSW2015, the difference in the cooperation rate ranges from -9 to $+13$ percentage points, with the corresponding values for NSW 2016 being -7 and $+16$ percentage points.

Not all of this variability is attributable to how skilful interviewers are in deploying incentives and simply reflects random variability in response propensities across interviewer assignments. We can get a better sense of the effect of interviewers on incentive effectiveness by taking the expected cooperation rate for an incentivized household using interviewer from the top and bottom deciles of the random coefficient variance, σ_{ij}^2 , while holding all other variables constant. For NSW2015, this shows that interviewers in the top performing decile achieve an expected cooperation rate of 67 percent for incentivized households compared with 64 percent for those in the bottom decile and compared with 68 percent for the median interviewer for nonincentivized households, a substantial difference. The corresponding figures for NSW2016 are 64 percent and 58 percent for the top and bottom deciles, respectively, and 65 percent for the median interviewer for nonincentivized households. There is no obvious relationship between the overall response rate and the effectiveness of

the incentive within interviewers, so we find no evidence that interviewers who are, on average, better at obtaining cooperation are also more effective in deploying the incentive.

Next, we turn to the same analysis of UKHLS-IP, which, as a household longitudinal survey, has a rather different design from the Welsh cross-sectional surveys, although we focus on wave one only. Table 3 presents the estimated coefficients, standard deviations, and corresponding 95 percent credible intervals. These are consistent with those presented in table 2; the fixed effect for the incentive predicting response is positive but nonsignificant, and the interviewer characteristics—age, gender, and experience—are all nonsignificant, as are the interactions between these variables and the incentive fixed effect.

Three of the area level variables are significantly associated with response; the higher the urbanicity and population mobility, the lower the level of survey response, while areas with a housing structure comprising more terraced housing and vacant properties have higher levels of response. Even after controlling for these differences in area composition, the random coefficient for the incentive is significant, with a variance of 0.13 (95 percent credible interval 0.03–0.35). This suggests that the between-interviewer variability in the effectiveness of the incentive is caused by interviewer behavior rather than by differences in the sorts of people they have been allocated to interview. The model DIC decreases by 3.10 with the inclusion of the random coefficient, so we also find evidence of a between-interviewer difference in the effectiveness of the incentive on this alternative measure of statistical significance.

As with the Welsh surveys, the covariance between the random intercept and random slope is positive but with a 95 percent credible interval that includes zero. We therefore also find no support from UKHLS-IP for the idea that interviewers who, on average, obtain higher response rates might also be more effective in their deployment of incentives.

Figure 2 plots the difference in the mean predicted rates of response for each interviewer derived as fitted values from the models in table 3. It shows a very similar pattern to what we saw in figure 1 for the Welsh surveys, with substantial between-interviewer variation in response rates between high and low incentive groups with a range of –21 to +18 percentage points. Visually, there is more evidence of a positive correlation between difference in response rates and the overall mean response rates for each interviewer, although this difference is not statistically significant.

5. DISCUSSION

John Wannamaker, the American department store magnate, once (apocryphally) observed that “half the money I spend on advertising is wasted; the trouble is I don’t know which half.” The same sentiment might also be applied to monetary incentives in surveys (Rossolatos, 2013), although in this context,

Table 3. Estimated Coefficients for the Final Model for Innovation Panel Response

Variable (reference category)		β	SD	95% credible interval	
Intercept	Category	0.383	0.647	−0.893	1.679
Incentive (£5 per adult)	£10 per adult	0.217	0.549	−0.471	1.767
Neighbourhood Characteristics					
Urbanicity		−0.213	0.078	−0.368	−0.061
Housing structure		0.302	0.077	0.153	0.455
Population Mobility		−0.269	0.094	−0.455	−0.089
Interviewer Characteristics					
Gender (Female)	Male	−0.477	0.304	−1.089	0.117
Age (less than 40 years)	41 to 50 years	0.550	0.654	−0.764	1.852
	50 to 60 years	0.730	0.512	−0.246	1.754
	> 60 years	0.251	0.675	−1.103	1.567
Experience (less than 2 yrs.)	3 to 6 years	−0.355	0.326	−1.010	0.258
	7 to 9 years	−0.658	0.413	−1.479	0.151
	>10 years	−0.869	0.570	−1.998	0.220
Cross-level interactions					
Incentive (£5 per adult)	£10 per adult *Male	0.140	0.143	−0.345	0.624
*Gender (Female)					
Incentive (£5 per adult)	£10 per adult	−0.024	0.611	−1.250	1.147
	*41 to 50 years				
	£10 per adult	−0.110	0.551	−1.194	0.951
	*50 to 60 years				
	£10 per adult	0.077	0.567	−1.040	1.165
Incentive (£5 per adult)	*> 60 years				
	£10 per adult	0.035	0.266	−0.496	0.552
	*Experience (less than 2 yrs.)				
	£10 per adult	0.366	0.327	−0.273	1.010
	*7 to 9 years				
Incentive (£5 per adult)	£10 per adult	0.198	0.428	−0.642	1.045
	*>10 years				
Random intercept					
$\sigma^2_{0jk} = var(\mu_{0jk})$		1.143	0.359	0.582	1.984
Random coefficient					
$\sigma^2_{1k} = var(\mu_{1k})$		0.126	0.085	0.028	0.349
Covariance of intercept & coefficient					
$\mu_{01} = cov(\mu_{0k}, \mu_{1k})$ (interviewer)		0.130	0.180	−0.274	0.349
DIC			2529.201		

UKHLS-IP, wave one, N = 2, 123.

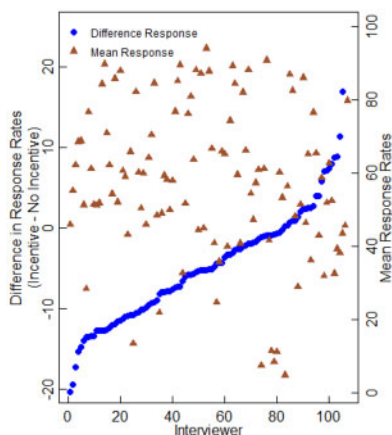


Figure 2. Difference in Predicted Rates of Response for Incentive and Nonincentive Households by Interviewer for UKHLS-IP.

considerably more than half of the money is wasted. This is because incentives generally add only a few percentage points or so to the headline response rate. It follows therefore that the majority of respondents in any survey using a monetary incentive would have agreed to provide an interview anyway. A small minority, however, are susceptible to being converted from refusal to interview with the provision of an incentive, and this in turn raises the possibility that interviewers might play an important role in determining the rate of such “conversions.” While there are other reasons for providing monetary incentives than boosting the response rate, this remains the primary rationale in most cases. It is therefore important to understand how best to maximize the effectiveness of monetary incentives in converting refusals to interviews. This is all the more pressing, given the likely need to place greater reliance on incentives to maintain response rates in the future.

Our findings show that across three different UK face-to-face surveys, interviewers vary significantly in how effective they are at using incentives to increase rates of cooperation. The effects we observe are substantively and statistically significant; our model estimates show that exchanging interviewers from the top to the bottom decile of interviewer performance would yield an expected 14 to 15 percentage point increase in the effect of the incentive relative to the control condition. We have speculated that this heterogeneity results from interviewer expectations and behavior, particularly the use of “tailoring” of doorstep interactions (Groves and Couper 1998) and greater confidence in the probability of obtaining an interview when an incentive is offered (Singer, Frankel, and Glassman 1983; Singer, Hoewyk, and Maher 2000). However, while the between-interviewer variability in the effectiveness of incentives was consistent across the three surveys, we found no significant predictor of this variance among the covariates considered: interviewer age, sex, and

experience. Nor was variability in incentive effectiveness related to the overall response rate an interviewer achieved. Therefore, the mechanisms underpinning this effect remain unclear.

Our focus in this article has been on the effect of incentives on cooperation because incentives seem likely to exert their primary influence on the cooperation decision. However, it is possible that they also have an effect on contact rates and other categories of nonresponse. We have therefore also carried out the analyses reported here with the dependent variable specified as response/nonresponse for the two surveys where full outcome codes were available before any reissues. The results are substantively identical to those reported here, so we find no evidence of a differential effect of interviewers on cooperation relative to total nonresponse.¹

Our findings have implications for survey practice. The approach we have implemented here to identify interviewer effectiveness in deploying incentives could be used as a way of identifying underperforming interviewers. This sort of monitoring is now implemented routinely in many large-scale survey operations, often in real-time, as a way of identifying interviewers who show signs of missing fieldwork targets (Kreuter 2013; Edwards, Maitland, and O'Connor 2017). It should be feasible to include “incentive performance” alongside other forms of paradata to raise flags against particular interviewers on this performance dimension, although how this would be adapted to designs in which all households are offered the same incentive would require further consideration.

Relatedly, the ability to identify interviewers at the top end of the performance distribution offers opportunities to better understand the types of strategies employed by more successful interviewers. Information on successful approaches to incentive use that are identified in this way could be integrated into sections of interviewer briefings that address doorstep approaches, both for generic and survey-specific training. Indeed, simply highlighting to interviewers that the way they administer incentives can have substantial effects on their response outcomes may, on its own, have some effect on their subsequent behavior.

While our methodological approach and findings represent an advance in our understanding of how interviewers and incentives interact to promote cooperation, this study is not without limitations, and these should be acknowledged. First, the surveys we have considered all use a relatively narrow range of incentive values which are administered to all households in the incentive condition. Caution should be exercised in generalizing to contexts where larger incentives are used or where incentives of varying values are targeted at different subgroups of the sample based on response propensities (Lavrakas, McPhee, and Jackson 2016). Our results also have little relevance to the use of incentives in online surveys, which comprise a large and growing proportion of total survey volume, both in the United Kingdom and internationally.

1. These analyses are available from the corresponding author upon request.

We were also able to link the sample file and response outcome data to a limited range of area and interviewer characteristics. It is possible that with stronger controls for differences between interviewers in the composition of their allocated addresses, the magnitude of the effects we have observed might be reduced. The paucity of interviewer characteristic data available to us, particularly the absence of variables measuring interviewer attitudes, beliefs, and behaviors, means that our ability to explain *why* some interviewers are more effective in deploying monetary incentives than others is weak. These limitations, we contend, represent potentially fruitful avenues for future research.

Appendix 1: Descriptions of Survey Sample Designs & Methodology

NATIONAL SURVEY FOR WALES FIELD TEST 2015 (NSW 2015)

The sample design of the NSW2015 used a stratified, single-stage, random selection of addresses across Wales drawn from the small user Postcode Address File (PAF). Adults 16 years old or over within each sampled household were interviewed face-to-face, and each interview lasted for an average of 25 minutes. When a household contained more than one adult, a single adult was randomly selected. The aim of the incentive experiment was to assess the extent to which response rates improved by offering respondents a £10 gift card upon completing an interview. The experimental group (N=2,965) received a £10 conditional incentive, and the second group received no incentive (N=2,830). The households which were randomly selected to be offered a conditional £10 received advance letters mentioning the incentive, while the other half of households received advance letters that contained no information about incentives. To ensure that any differences in response rates between respondents who were offered £10 and those offered no incentive are not attributed to any interviewer abilities, addresses that were offered incentives were randomly allocated within each interviewer assignment. The survey was implemented by a team of 86 interviewers with the number of households interviewed by each interviewer ranging between 14 and 134. Further details on the NSW2015 sample design can be found in [Hanson, Sullivan, and McGowan \(2015\)](#).

NATIONAL SURVEY FOR WALES INCENTIVE EXPERIMENT 2016 (NSW 2016)

The Welsh government commissioned the office for National Office of National Statistics (ONS) to conduct the National Survey for Wales 2016 (NSW 2016) incentive experiment between July and October 2016. The

sample was drawn from the Postcode Address File (PAF). The stratification was done by local authority (LA), using an allocation designed to ensure a minimum effective sample size was achieved in each LA based on estimated response rate. Further details on the sample design can be found in [Aumeyr et al. \(2017\)](#). Half of the addresses in each odd numbered quota² were offered a £5 incentive conditional on participation ($N = 3,604$), and addresses with even quota number were offered no incentive ($N = 3,467$). The incentive experiment ran from July to October 2016. Originally, it was intended to run the experiment until December 2016, but it was terminated at the end of October 2016 because both experimental and control groups experienced lower response rates at 55 percent and 54 percent, respectively, which were lower than expected. With an aim of boosting response rates, a new £10 incentive conditional on participation was introduced to the full sample in November 2016. This study only considers the experiment sample size from July to October 2016 that consists of 7,071 households across the two conditions. There were 85 interviewers working on the survey with the minimum and maximum number of interviews per interviewer ranging between one and 219. Sociodemographic characteristics of ten (12 percent) interviewers who conducted interviews on 249 (3.5 percent) households were missing because they did not provide consent. The final analysis sample had 6,122 households after excluding 742 (10.5 percent) ineligible households and those interviewed by interviewers with missing sociodemographic characteristics.

UK HOUSEHOLD LONGITUDINAL SURVEY INNOVATION PANEL WAVE ONE (UKHLS-IP)

The sample for wave one of the UK Household Longitudinal Survey Innovation Panel (UKHLS-IP) was clustered and stratified, consisting of 2,786 addresses from 120 primary sampling units (PSUs) from the Postcode Address File (PAF). The incentive experiment was comprised of three conditions, with each condition receiving a different conditional incentive: Group one was offered £5 per adult, group two was offered £10 per adult, and group three was offered £5 per adult, rising to £10 per adult if all adults in the household completed interviews. Single person households randomly assigned to group three received £5 initially that increased to £10 if they participated. For the purposes of our analysis, groups two and three are combined. Note that all households were also sent an unconditional £5 incentive with the advance letter. There were twenty-seven households in the UKHLS-IP that did not successfully merge with interviewer data due to lack of common unique identifiers. The neighborhood characteristic variables are drawn from the

2. Each quota contained between twenty and thirty addresses on average.

census and were available for England only. This resulted in the exclusion of 342 households (12.3 percent) from Wales and Scotland. In addition, thirty-one (1.1 percent) households in five MSOAs in England did not successfully merge with Innovation Panel data due to lack of common unique identification codes. Therefore, the final analysis sample contained 2,123 households after excluding 263 (9.4 percent) ineligible households. The number of interviewers working on the UKHLS-IP was 107, with the number of households interviewed by each interviewer ranging between two and fifty. Further details about the UKHLS-IP can be found in [Boreham and Constantine \(2008\)](#)

References

- AAPOR (2016), *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 9th edition. AAPOR. The American Association for Public Opinion Research.
- Aumery, M., Z. Brown, R. Doherty, A. Fallows, T. Pegg, R. Perez-Dominguez, R. Sanderson, I. Shemwell (2017), *National Survey for Wales 2016-17 Technical Report*. Cardiff. Available at <http://www.gov.wales/nationalsurvey>. Retrieved February 13 2018.
- Blom, A. G., E. de Leeuw, and J. J. Hox (2011), "Interviewer Effects on Nonresponse in the European Social Survey," *Journal of Official Statistics*, 27, 359–377.
- Blom, A. G., and J. M. Korbmacher (2013), "Measuring Interviewer Characteristics Pertinent to Social Surveys: A Conceptual Framework," *Survey Methods: Insights from the Field*, 1–16. Available at <https://doi.org/10.13094/SMIF-2013-00001>.
- Boreham, A. R., and R. Constantine (2008), *Understanding Society Innovation Panel Wave 1 Technical Report*, (7083), 1–14.
- Brick, M. J., J. M. Montaquila, M. C. Hagedorn, S. B. Roth, and C. Chapman (2005), "Implications for RDD Design from an Incentive Experiment," *Journal of Official Statistics*, 21, 571–589.
- Brick, M. J., and D. K. Williams (2013), "Explaining Rising Nonresponse Rates in Cross-Sectional Surveys," *The ANNALS of the American Academy of Political and Social Science*, 645, 36–59.
- Browne, W. J., M. Kelly, C. Charlton, and R. Pillinger (2016), *MCMC Estimation in MLwiN: Version 2.36*.
- Campanelli, P., and C. O'Muirheartaigh (1999), "Interviewers, Interviewer Continuity, and Panel Survey Nonresponse I," *Quality & Quantity*, 33, 59–76.
- Campanelli P., P. Sturgis, and S. Purdon (1997), *Can You Hear Me Knocking: An Investigation into the Impact of Interviewers on Survey Response Rates*, London: SCPR Survey Methods Centre.
- Cantor, D., B.C. O'Hare, and K.S. O'Connor (2008). "The Use of Monetary Incentives to Reduce Nonresponse in Random Digit Dial Telephone Surveys," in *Advances in Telephone Survey Methodology*, eds. J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japec, P. J. Lavrakas, M.W. Link, and R. L. Sangster, pp. 471–498. Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9780470173404.ch22>
- Church, A. H. (1993), "Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis," *Public Opinion Quarterly*, 57, 62–79.
- Couper, M. P. (2013), "Is the Sky Falling? New Technology, Changing Media, and the Future of Surveys," *Survey Research Method*, 7, 145–156.
- Couper, M. P., and J. Schlegel (1998), Evaluating the NHIS CAPI Instrument Using Trace Files. In *Annual Meeting of the American Association for Public Opinion Research*. St. Louis. Available at https://ww2.amstat.org/sections/srms/Proceedings/papers/1997_142.pdf.
- Dillman, D. A., J. D. Smyth, and L. M. Christian (2009), *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method* (3rd ed.), New York: John Wiley & Sons.

- Durrant, G. B., R. M. Groves, L. Staetsky, and F. Steele (2010), "Effects of Interviewer Attitudes and Behaviors on Refusal in Household Surveys," *Public Opinion Quarterly*, 74, 1–36.
- Durrant, G. B., and F. Steele (2009), "Multilevel Modelling of Refusal and Non-Contact in Household Surveys: Evidence from Six UK Government Surveys," *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 172, 361–381.
- Edwards, B., A. Maitland, and K.S. O'Connor (2017), "Measurement Error in Survey Operations Management: Detection, Quantification, Visualization, and Reduction," in *Total Survey Error in Practice*, eds. P. Biemer, E. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. Lyberg, N. C. Tucker, and B. T. West, Hoboken, NJ: John Wiley & Sons.
- Fearn, T., A. Gelman, J. B. Carlin, H. S. Stern, D. B. Rubin, and T. Fearn (2004), "Bayesian Data Analysis," *Biometrics*, 52, 696.
- Gelman, A., M. Stevens, and V. Chan (2002), "Regression Modeling and Meta-Analysis for Decision Making: A Cost-Benefit Analysis of Incentives in Telephone Surveys," *Journal of Business & Economic Statistics*, 21, 213–225.
- Goldstein, H. (2010), *Multilevel statistical models* (4th ed.), Chichester: John Wiley & Sons.
- Groves, R. M., and M. P. Couper (1996), "Contact Level Influences on Cooperation in Face to Face Surveys.pdf," *Journal of Official Statistics*, (12), 63–83.
- . (1998), *Nonresponse in Household Interview Surveys*. New York: John Wiley and Sons.
- Groves, R. M., and S. G. Heeringa (2006), "Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs," *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 169, 439–457.
- Groves, R. M., E. Singer, and A. D. Corning (2000), "Leverage-Saliency Theory of Survey Participation Description and an Illustration," *Public Opinion Quarterly*, 64, 299–308.
- Hansen, K. M. (2006), "The Effects of Incentives, Interview Length, and Interviewer Characteristics on Response Rates in a CATI-Study," *International Journal of Public Opinion Research*, 19, 112–121.
- Hanson, T. E., S. Sullivan, and C. McGowan (2015), *National Survey for Wales 2014-15 Technical Report*. Cardiff.
- Hox, J. J., and E. de Leeuw (2002), "The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: An International Comparison," in *Survey Nonresponse*, eds. R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little, pp. 103–120, New York: Wiley.
- Kreuter, F. (2013), *Improving Surveys with Paradata: Analytic Uses of Process Information*, ed. F. Kreuter, Hoboken, NJ: John Wiley & Sons.
- Lavrakas, P. J. (2008), *Encyclopedia of Survey Research Methods*. Thousand Oaks, CA: SAGE Publications.
- Lavrakas, P. J., C. McPhee, and M. Jackson (2016), "Conceptual Background on Response Propensity Modeling (RPM) for Allocating Differential Survey Incentives: JSM 2016—Survey Research Methods Section 3280 Purpose, Rationale, and Operationalization," Paper Presented at the 71st Annual Conference of the American Association for Public Opinion Research, Austin, TX.
- Lemay, M., and C. Durand (2002), "The Effects of Interviewer Attitude on Survey Cooperation," *Bulletin de Methodologie Sociologique*, 76, 27–44.
- Lynn, P. (2001), "The Impact of Incentives on Response Rates to Personal Interview Surveys: Role and Perceptions of Interviewers," *International Journal of Public Opinion Research*, 13, 326–336.
- Mercer, A., A. Caporaso, D. Cantor, and R. Townsend (2015), "How Much Gets You How Much? Monetary Incentives and Response Rates in Household Surveys," *Public Opinion Quarterly*, 79, 105–129.
- Meyer, B. D., W. K. C. Mok, and J. X. Sullivan (2015), "Household Surveys in Crisis," *Journal of Economic Perspectives*, 29, 199–226.
- Miller, P. V. (2017), "Is There a Future for Surveys?" *Public Opinion Quarterly*, 81, 205–212.
- Morton-Williams, J. (1993), *Interviewer Approaches*, Aldershot: Dartmouth Publishing Company Limited.

- O'Muircheartaigh, C., and P. Campanelli (1998), "The Relative Impact of Interviewer Effects and Sample Design Effects on Survey Precision," *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 161(1), 63–77.
- Purdon, S., P. Campanelli, and P. Sturgis (1999), "Interviewers' Calling Strategies on Face-to-Face Interview Surveys," *Journal of Official Statistics*, 15, 199–216.
- Rasbash, J., F. Steele, W. J. Browne, H. Goldstein, C. Charlton, and W. J. Browne (2012), *A User's Guide to MLwiN*, v2.26. Centre for Multilevel Modelling, United Kingdom: University of Bristol.
- Rees, P. H. (1971), "Factorial Ecology: An Extended Definition, Survey and Critique of the Field," *Economic Geography*, 47, 220–233.
- Rossolatos, G. (2013). "I know half the money I spend on advertising is wasted, but I do not know which half (J. Wannamaker)": Semiotic answers to perennial branding troubles. *Social Semiotics*, 23(4), 545–560. <https://doi.org/10.1080/10350330.2013.79900>
- Singer, E. (2002), "The Use of Incentives to Reduce Nonresponse in Household Surveys," in *Survey Nonresponse*, eds. R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little, pp. 163–177, New York: Wiley.
- Singer, E., M. R. Frankel, and M. B. Glassman (1983), "The Effect of Interviewer Characteristics and Expectations on Response," *Public Opinion Quarterly*, 47, 68–83.
- Singer, E., R. M. Groves, and A. D. Corning (1999), "Differential Incentives: Beliefs about Practices, Perceptions of Equity, and Effects on Survey Participation," *Public Opinion Quarterly*, 63, 251–260.
- Singer, E., J. V. Hoewyk, N. Gebler, T. Raghunathan, and K. McGonagle (1999), "The Effect of Incentives on Response Rates in Interviewer-Mediated Surveys," *Journal of Official Statistics*, 15, 217–231.
- Singer, E., J. V. Hoewyk, and M. P. Maher (2000), "Experiments with Incentives in Telephone Surveys," *Public Opinion Quarterly*, 64, 171–188.
- Singer, E., and L. Kohnke-Aguirre (1979), "Interviewer Expectation Effects: A Replication and Extension," *Public Opinion Quarterly*, 43, 245–260.
- Singer, E., and C. Ye (2013), "The Use and Effects of Incentives in Surveys," *The ANNALS of the American Academy of Political and Social Science*, 645, 112–141.
- Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde (2002), "Bayesian Measures of Model Complexity And Fit," *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, 64, 583–639.
- Stoop, I. A. L. (2005), *The Hunt for the Last Respondent*, The Hague: Social and Cultural Planning Office (SCP).
- West, B. T., and A. G. Blom (2017), "Explaining Interviewer Effects: A Research Synthesis," *Journal of Survey Statistics and Methodology*, 5, 175–211.